

# Reforming the System of Review by Administrative Law Judges in Disability Insurance

---

Mark J. Warshawsky and Ross A. Marchand

*September 2015*

MERCATUS WORKING PAPER



3434 Washington Blvd., 4th Floor, Arlington, Virginia 22201  
[www.mercatus.org](http://www.mercatus.org)

*Mark J. Warshawsky and Ross A. Marchand. "Reforming the System of Review by Administrative Law Judges in Disability Insurance." Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, September 2015.*

## **Abstract**

In this study we update and deepen Mark Warshawsky's prior examination of the administrative law judge system for the Social Security disability program. We add reviews of recent comprehensive reports from academic and government sources and undertake our own analysis of judicial decisions over a longer time period, using recent data. These varied examinations each reach similar conclusions to one another. Our particular estimate is that failings by the administrative law judge system have led, on net, to more than \$72 billion being paid out to disability claimants over their lifetimes through unwarranted benefit awards over the period of 2005 through 2014. Recent public scrutiny and administrative changes have curbed some of the worst excesses, but serious problems remain and could worsen again when attention moves elsewhere or when political pressure builds to reduce the claims backlog. Therefore permanent administrative reform is needed to lock in the recent changes and to build on them to enhance their positive effect.

*JEL* codes: H55, I38, J18

Keywords: Social Security, disability, administrative law, adjudication of benefit claims

## **Author Affiliation and Contact Information**

Mark J. Warshawsky, PhD  
Senior Research Fellow  
Mercatus Center at George Mason University  
Visiting Fellow  
MIT Center for Finance and Policy  
mwarshawsky@mercatus.gmu.edu

Ross A. Marchand  
MA Fellow  
Mercatus Center at George Mason University  
George Mason University Department of Economics  
rmarcha2@gmu.edu

*All studies in the Mercatus Working Paper series have followed a rigorous process of academic evaluation, including (except where otherwise noted) at least one double-blind peer review. Working Papers present an author's provisional findings, which, upon further consideration and revision, are likely to be republished in an academic journal. The opinions expressed in Mercatus Working Papers are the authors' and do not represent official positions of the Mercatus Center or George Mason University.*

## Reforming the System of Review by Administrative Law Judges in Disability Insurance

Mark J. Warshawsky and Ross A. Marchand

The Social Security Disability Insurance and Supplemental Security Income programs have many financial and structural problems. In a recent editorial, we detailed problems with the system of benefit claims appeals presided over by administrative law judges, and we proposed several solutions.<sup>1</sup> Because the editorial, based on an earlier study by Mark Warshawsky,<sup>2</sup> was published in the *Wall Street Journal*, it garnered attention from the public, the media, and others. We therefore thought it worthwhile to continue, update, and deepen our examination of the topic, adding reviews of some recent comprehensive reports—econometric, analytical, and case studies—from academic and government sources. Additionally, we undertook our own empirical analysis of judicial decisions over a longer period that includes more recent data.

It is difficult to come to a definitive conclusion based on any one study, report, or analysis, but when varied sources reach similar conclusions through different methodologies and approaches, they build a compelling case. Here we find such a case, in which serious failings by the administrative law judge system—at a time of large claims backlogs—have led, on net, to large losses for taxpayers. We estimate that more than \$72 billion will be paid to claimants over their lifetimes through likely unwarranted disability benefit awards given by administrative law judges over the 10-year period of 2005 through 2014. Recent public scrutiny and administrative changes have curbed some of the worst excesses, but serious problems with the benefit claims appeals process remain. Moreover, the problems' original magnitude could easily return when

---

<sup>1</sup> Mark Warshawsky and Ross Marchand, “Disability Claim Denied? Find the Right Judge,” *Wall Street Journal*, March 8, 2015.

<sup>2</sup> Mark Warshawsky, “Administrative Problems with Social Security Disability Programs: Some Solutions,” *Bloomberg Pension and Benefits Daily*, April 2, 2012.

public and management attention move elsewhere or when political pressure builds again to reduce the claims backlog. Therefore, serious permanent administrative reform is needed to lock in the recent changes and to build on them to enhance their positive effect. The appropriate time to do so is now, as the disability insurance program heads to insolvency in two years, as other reforms are needed and hopefully legislated to tighten eligibility,<sup>3</sup> and as there are calls for transfers of more taxpayer resources to the program.

## **Background**

The Social Security Administration (SSA) manages two large federal disability programs: Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI). As of 2014, about 19.4 million individuals receive about \$200 billion annually in benefits through these two programs. In addition to the cash benefits, individuals enrolled in SSDI for two years are automatically enrolled in Medicare, costing taxpayers about \$80 billion a year for SSDI beneficiaries. SSI recipients are eligible for Medicaid immediately.

When an individual applies for disability benefits, the case is initially decided by state employee examiners in Disability Determination Services (DDS). There is also an automatic preeffectuation review—an internal review of a case decision before it is finalized and communicated to the claimant—of 50 percent of DDS allowance decisions. In 40 states and most of California, an applicant who is denied benefits may appeal to a different reviewer in the same office. The SSA, which oversees each state’s DDS, claims that there are few errors in the original adjudication of these decisions. If the second reviewer denies benefits, the applicant may

---

<sup>3</sup> See Mark Warshawsky and Ross Marchand, “Modernizing the SSDI Eligibility Criteria: A Reform Proposal That Eliminates the Outdated Medical-Vocational Grid” (Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, April 2015).

appeal to an administrative law judge (ALJ).<sup>4</sup> If the ALJ then awards disability benefits, the decision is final because the government does not appeal it. But if the ALJ denies benefits, the individual may appeal further, to the SSA Appeals Council and, as a civil case, to the several levels of the federal courts. In total, there are at least five levels of review for a disability benefits applicant. Any error by a state adjudicator or an ALJ against an applicant is fixable, whereas an error by them against the taxpayer is not.

According to economists David Autor and Mark Duggan, the average lifetime disability benefit, including cash and the value of health care, is about \$300,000.<sup>5</sup> During the 1980s, the disability programs experienced loosened eligibility requirements and increased benefits. Autor and Duggan find that the result was a rise in the proportion of men who report work limitations and a concomitant decrease in the employment of men with disabilities.<sup>6</sup> This trend has continued since their study. General survey indicators of disability rates in the working-age population are flat or declining.<sup>7</sup> SSA economists T. Von Wachter, Jay Song, and Joyce Manchester find that rejected applicants who are young and who have low-mortality reported disabilities show a relatively strong attachment to the labor force.<sup>8</sup>

---

<sup>4</sup> The SSA employs almost 1,400 ALJs to adjudicate about 700,000 cases a year. These independent judges effectively have lifetime tenure and are to conduct impartial “de novo” hearings and make decisions on appealed agency determinations. ALJs can only be removed from office through a lengthy and costly process conducted by the US Merit Systems Protection Board; such removals are rare. The 1946 Administrative Procedure Act (APA) governs administrative adjudication. Under the Constitution, ALJs are Article I judges, which the Supreme Court has recognized as “functionally comparable” to Article III trial judges. ALJs are not supervised by anyone engaged in agency investigative or prosecutorial functions, are immune from liability for judicial acts, and are exempt from performance ratings, evaluation, and receipt of bonuses. Their salaries are not set by the agency. For this and further information, see Judge Thomas P. McCarthy, “Respect Administrative Law Judges,” letter to the *Wall Street Journal*, August 28, 2014.

<sup>5</sup> David Autor and Mark Duggan, “Supporting Work: A Proposal for Modernizing the U.S. Disability Insurance System” (Center for American Progress Paper, Center for American Progress and the Hamilton Project, Washington, DC, 2010).

<sup>6</sup> David Autor and Mark Duggan, “The Rise in the Disability Rolls and the Decline in Unemployment,” *Quarterly Journal of Economics* 118 (2003): 157–206.

<sup>7</sup> Warshawsky and Marchand, “Modernizing the SSDI Eligibility Criteria.”

<sup>8</sup> T. Von Wachter, Jay Song, and Joyce Manchester, “Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program,” *American Economic Review* 101 (2011): 3308–29.

According to George Washington University law professor Richard Pierce, reversals of SSA denials by ALJs have increased significantly from 1970 to the present, primarily because ALJs have granted benefits to many applicants with less severe mental illness and pain than ALJs considered sufficient to qualify for disability benefits in the past.<sup>9</sup> About 60 to 70 percent of ALJ rulings in disability benefit appeals, on average, have been in the claimant's favor in recent years. ALJs have a greater incentive to award benefits than to deny them because denials are subject to judicial appeal, and because denials must be fully documented, which takes longer, whereas decisions and drafting of approvals are typically quick.<sup>10</sup>

The approval rate has fallen recently, to about 54 percent. This decline is perhaps a reaction by the agency and the ALJs to negative publicity arising from investigative reports from the *Wall Street Journal* about ALJs who ruled for claimants in virtually all their cases.<sup>11</sup> The reports highlight Judge David Daugherty from the Office of Disability Adjudication and Review in Huntington, West Virginia. He heard thousands of cases over the years and uniformly granted favorable judgements. Many of these cases were represented by one particular law firm. There have also been strong congressional inquiries, a series of hearings, and administrative reforms by former Social Security commissioner Michael Astrue. In addition, a lagged business-cycle effect is likely a factor in the recent decline because of an increased number of claims motivated mainly by high unemployment during the Great Recession rather than by personal disability and therefore clearly subject to denial. Former commissioner Astrue's administrative reforms will be described in more detail later.

---

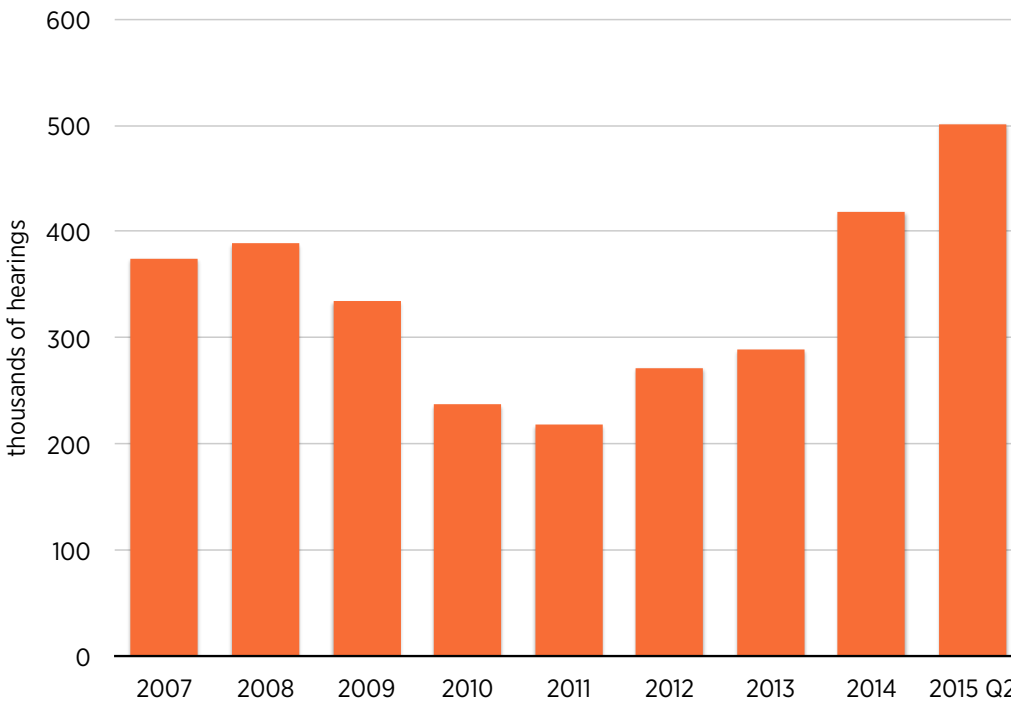
<sup>9</sup> Richard Pierce, "What Should We Do about Administrative Law Judge Disability Decision Making?" (Public Law and Legal Theory Paper 573, George Washington University Law School, Washington, DC, 2011).

<sup>10</sup> Randy Frye and Marilyn Zahm, interview by Steve Kroft, "Easier to Approve a Disability Case Than Deny It?," *60 Minutes*, CBS, October 6, 2013; email from the SSA's Office of Inspector General (OIG) to staff of the House Committee on Oversight and Government Reform.

<sup>11</sup> See, for example, Damian Paletta, "Disability-Claim Judge Has Trouble Saying 'No,'" *Wall Street Journal*, May 19, 2011.

Another piece of relevant historical background is the backlog of disability claims hearings pending for more than 270 days. As figure 1 shows, the backlog had reached more than 400,000 cases in 2008. There was strong political pressure on the SSA to reduce it, which the agency did over the next few years. Then the backlog began to build again in 2012, and in 2014, it was again more than 400,000; by the second quarter of fiscal year 2015, the backlog reached 500,000 cases.

**Figure 1. Number of Hearings Pending for More Than 270 Days, FY 2007–Q2 FY 2015**



Source: data from the Social Security Administration.

The SSDI trust fund is scheduled to go bankrupt by late 2016, at which time benefits will have to be cut by 19 percent to all SSDI beneficiaries so the program will remain solvent and keep within the program’s cash flows.<sup>12</sup>

<sup>12</sup> SSA Board of Trustees, “The 2014 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds,” July 2014, House Document 113–39.

## **An Important Econometric Study**

Economists Robert Nakosteen and Michael Zimmer conducted an econometric analysis of decisions by administrative law judges based on data from cases that came before about 1,500 judges for the fiscal years ended September 2010 through September 2012.<sup>13</sup> They examined the data for patterns in both the approval rate and volume of decisions rendered. They used information about the judges' gender and number of years of judicial experience, as well as the state unemployment rate, an indicator of political makeup for the state in which the judge presided, and the number of decisions each judge rendered.

Nakosteen and Zimmer show basic statistical evidence of an upward drift in approvals by ALJs as a function of decision volume; judges deciding many cases trend toward leniency. They find that mean judicial experience is 31 years and that experience is positively related to approval rates. They also find that judges tend to be lenient in environments of relatively high joblessness in the state and in the presence of a Democratic governor. There is no difference in approval rates between genders among the judges. Nakosteen and Zimmer also find that the general disposition toward leniency declined from 2010 through 2012, and although they are not certain of the cause, they emphasize the impact of negative publicity from a series of *Wall Street Journal* investigative articles. Finally, they find econometric evidence that there is a tendency for lenient judges to take large caseloads and this tendency has grown more pronounced over time.

## **Analytical Studies**

One criticism of some studies examining ALJs is that they focus on lenient judges. A 2014 US Office of Inspector General (OIG) report examining low-approval judges, however, concludes

---

<sup>13</sup> Robert Nakosteen and Michael Zimmer, "Approval of Social Security Disability Appeals: Analysis of Judges' Decisions," *Applied Economics* 46, no. 23 (2014): 2783–91.



that the remand and reversal rate (used by the OIG as a quality performance proxy) for these low-approval ALJs is, on average, about the same as that of the general ALJ population.<sup>14</sup> The report also finds that 8 of the 12 lowest-allowance judges “decided fewer cases than the average of their peers.”<sup>15</sup> This finding is telling, given that the SSA regards ALJ decision count as a strong inverse predictor of decisional quality. After examining data from the Office of Disability Adjudication and Review (ODAR), the SSA concluded in a 2012 internal memo that there exists a “strong relationship between production levels and decisional quality on allowances. As ALJ production increases, the general trend for decisional quality on allowances falls.”<sup>16</sup> Thus, we can say from the relationship between decision volume and allowance rate established in the Nakosteen and Zimmer econometric analysis that high-allowance judges have lower decisional quality than low-allowance judges. Legal academics Harold Krent and Scott Morris find that the number of years spent by ALJs in the top 1 percent of allowance rates strongly predicts dispositional volume.<sup>17</sup> This conclusion, coupled with the OIG’s findings on low-allowance ALJ dispositional volume, suggests a large “quality gap” between high and low outlier judges.

In another report, the OIG uses quality review data to compare the accuracy of affirmative ALJ decisions with rejection decisions over the 2009–2010 period using randomly selected cases.<sup>18</sup> Examining 1,022 denials and allowances, the report finds that the approval decisions are 40 percent more likely to garner “disagreeable” ratings by the ODAR’s

---

<sup>14</sup> SSA Office of the Inspector General, “Subsequent Appellate Actions on Denials Issued by Low-Allowance Administrative Law Judges,” July 2014, A-12-13-13084.

<sup>15</sup> *Ibid.*

<sup>16</sup> SSA, “Memo on Production Levels and Decision Quality,” September 2012.

<sup>17</sup> Harold J. Krent and Scott B. Morris, “Inconsistency and Angst in District Court Resolution of Social Security Disability Appeals,” Chicago-Kent College of Law Research Paper No. 2014-30 (Chicago: Chicago-Kent College of Law, 2014).

<sup>18</sup> SSA Office of the Inspector General, “The Social Security Administration’s Review of Administrative Law Judges’ Decisions,” March 2012, A-07-12-21234.

posteffectuation review process. This finding is hardly surprising, given that hearing-stage applicants have already been rejected at the first two rounds of application by trained examiners at DDS. The SSA conducts preeffectuation quality reviews of randomly selected DDS decisions to determine whether these examiners reached a sound decision. In FY 2010–2014, the DDS accuracy rate consistently remained higher than 97 percent. By stating that DDS workers come to nearly perfect decisions during the first two stages of determination, the SSA is implying that subsequent appeals to ALJs should have low allowance rates. The existence and persistence of high-approval outlier judges, then, is a cause for concern.

Finally, an OIG report on judicial motivation finds that ALJs are deciding cases on considerations other than what the law and regulations allow, with the ALJs being inappropriately influenced by the unemployment conditions in the local economy or by personal considerations, such as their past occupations or political views.<sup>19</sup>

### **Case Studies**

In 2014, the House Committee on Oversight and Government Reform released an indictment of “rubber-stamping disability judges.”<sup>20</sup> The committee relied on SSA internal investigations of ALJs with high disposition counts and award rates. One was Charles Bridges, the hearing office chief ALJ for Harrisburg, Pennsylvania. Despite awarding benefits without holding hearings in 7,000 cases and being repeatedly criticized by the SSA for poor decisional quality, at least through 2014, he still enjoyed a full caseload.<sup>21</sup> ALJ Harry Taylor, repeatedly accused by

---

<sup>19</sup> SSA Office of the Inspector General, “Oversight of Administrative Law Judge Workload Trends,” February 2012, A-12-11-01138.

<sup>20</sup> House Committee on Oversight and Government Reform, “Systemic Waste and Abuse at the Social Security Administration: How Rubber Stamping Disability Judges Cost Hundreds of Billions of Taxpayer Dollars,” June 10, 2014.

<sup>21</sup> *Ibid.*, 5–6.

colleagues of conducting “sloppy work” and sleeping on the job, decided nearly 70 percent of cases without a hearing and denied awards to only 6 percent of claimants. Despite being handed a 14-day suspension for misconduct by the SSA and being recommended for another in 2013, at least through 2014, he still served at the bench.<sup>22</sup>

A second committee report on the matter recommended capping the number of annual dispositions to 600,<sup>23</sup> consistent with SSA research showing decisional quality delay for ALJs taking more than 617 dispositions in a given year. The House report also calls for a prioritization of resources devoted to ALJ decisional review; it states that hiring more ALJs should be put on hold until review capacity increases fivefold. More boldly, the committee concludes that judges found to be incorrectly applying disability law should be removed, and only reinstated upon completion of an observed, compliant trial period. Critically, the committee proposal would make high allowance rates sufficient to warrant further investigation of an ALJ, which could result in dismissal.<sup>24</sup>

We see two problems with these reports’ recommendations. First, it is challenging to remove an ALJ for performance through the existing administrative process. Commissioner Astrue was only able to remove three or four judges, despite considerable efforts, and these cases were blatant violations.<sup>25</sup> Second, the SSA has claimed, with some justification, that legally it cannot focus on judges with high approval rates, per se. That being said, more resources should be devoted to the review and analysis of ALJ decisions.

---

<sup>22</sup> Ibid., 6–7.

<sup>23</sup> House Committee on Oversight and Government Reform, “Misplaced Priorities: How the Social Security Administration Sacrificed Quality for Quantity in the Disability Determination Process,” December 2014.

<sup>24</sup> Ibid., 49–51.

<sup>25</sup> House Committee on Ways and Means, Subcommittee on Social Security, and House Committee on the Judiciary, Subcommittee on the Courts, *Statement of Michael J. Astrue Commissioner, Social Security Administration*, 112th Cong., 1st sess., July 11, 2011.

## **Our Extended Empirical Analysis**

As in Warshawsky's 2012 study,<sup>26</sup> in our new study, we calculate the net cost to taxpayers, per year and in total, of presumptively wrong decisions. Here we assume a cost of \$300,000 per case. We consider a decision as presumptively wrong based on a combination of fixed approval or denial rate numbers and twice the standard deviations in that year of all judges' decisions on both the approval and denial sides of the adjudication distribution. More specifically, we consider as presumptively in error the decisions of all judges with approval rates higher than 80 percent or beyond two standard deviations of the median on the right side of the distribution, as well as all judges with approval rates below 20 percent or beyond two standard deviations on the left side. The addition of outliers in terms of standard deviations (a relative measure) and not just fixed numbers accounts for the natural movement of the average ALJ's performance owing to the business-cycle effect. Two standard deviations represents behavior on the edges. The fixed numbers of 80 percent and 20 percent are equidistant from a 50 percent approval rate, a bit lower than the average of the last decades, but consistent with 80 percent being indicated by a prominent regional chief administrative law judge, Jasper Bede, as a "red flag" for problems in adjudication.<sup>27</sup>

Table 1 shows the average approval rates of ALJs and the standard deviations. Table 2 shows the proportion of ALJs with approval rates that are two standard deviations above and below the mean, and table 3 shows the proportion of ALJs with approval rates greater than 80 percent and lower than 20 percent, from 2005 through 2014. Note that approval rates are computed after taking out case dismissals, which are usually administrative actions and do not

---

<sup>26</sup> Warshawsky, "Administrative Problems."

<sup>27</sup> House Committee on Oversight and Government Reform, Subcommittee on Energy Policy, Health Care, and Entitlements, *Continuing Oversight of the Social Security Administration's Management of Federal Disability Programs*, 113th Cong., 1st sess., November 19, 2013.

indicate substantive adjudication. We consider both standard deviations (a relative measure) and percentage rates (a fixed measure), in both directions, as indicators of worrisome outliers. One standard deviation contains 66 percent of the ALJs around the mean. Figure 2 shows the distribution of judges by allowance rate from 2005 through 2013.

**Table 1. Average Allowance Rates of Administrative Law Judges, 2005–2014**

Year	Average approval rate (%)	Standard deviation (%)
2005	71.3	16.0
2006	70.6	16.3
2007	70.6	16.5
2008	69.6	16.5
2009	70.5	17.7
2010	67.1	16.3
2011	62.2	16.7
2012	57.2	16.2
2013	55.0	15.5
2014	53.7	15.1

Source: Authors' calculations based on data from the Social Security Administration.

**Table 2. Percentage of Administrative Law Judges with Approval Rates Two Standard Deviations above and below the Mean, 2005–2014**

Year	Judges 2+ SD above mean (%)	Judges 2+ SD below mean (%)
2005	0.00	3.20
2006	0.00	3.30
2007	0.00	3.50
2008	0.09	3.70
2009	0.00	3.70
2010	1.50	2.80
2011	2.20	2.50
2012	2.10	2.10
2013	2.40	1.70
2014	2.90	2.50

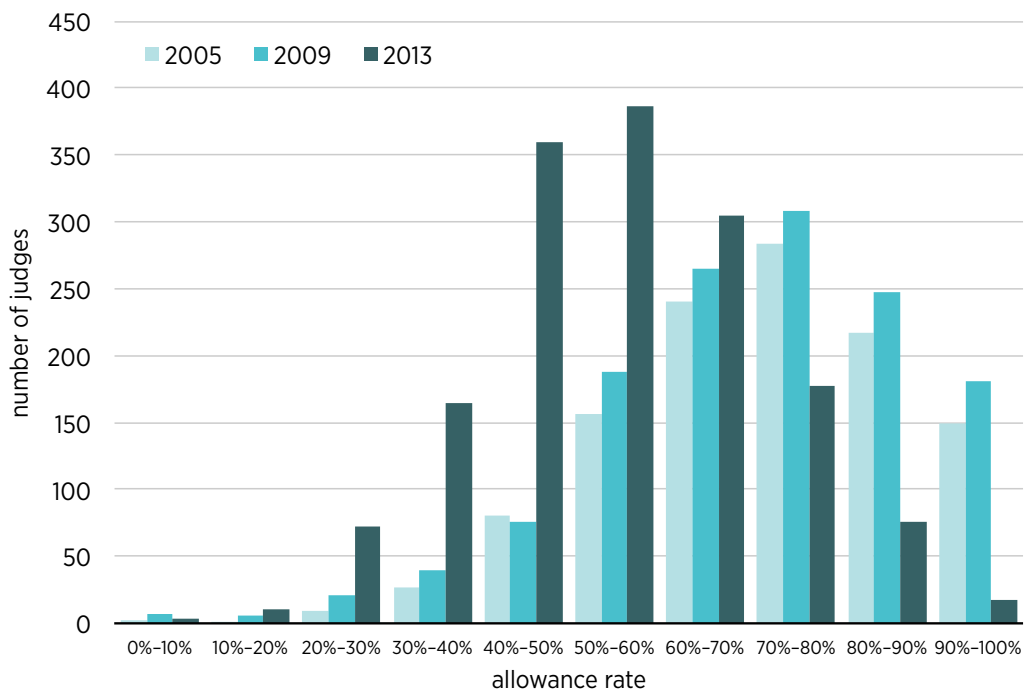
Source: Authors' calculations based on data from the Social Security Administration.

**Table 3. Percentage of Administrative Law Judges with Approval Rates of 80 Percent or Higher and 20 Percent or Lower, 2005–2014**

Year	% of judges with 80%+ approval	% of judges with 20%– approval
2005	31.4	0.3
2006	31.3	0.6
2007	31.4	0.8
2008	28.5	0.8
2009	32.0	1.0
2010	23.3	0.5
2011	16.1	0.7
2012	8.6	1.2
2013	5.9	0.9
2014	4.9	1.0

Source: Authors’ calculations based on data from the Social Security Administration.

**Figure 2. Distribution of Judges by Allowance Rate, 2005–2013**



Source: Authors’ calculations and analysis, based on data from the Social Security Administration.

Several things are evident from these tables and from figure 2. The average allowance rate dropped from 71 percent in 2005 to 54 percent in 2014, while the standard deviation

dropped from 18 percent in 2009 to 15 percent in 2014. The proportion of ALJs with allowance rates more than two standard deviations above the mean increased from zero to about 3 percent from 2005 through 2014, while the proportion with rates more than two standard deviations below the mean has bounced around somewhat, but overall has dropped from a high of almost 4 percent to around 2 percent recently. By this relative measure of “outlierness,” the number of generous ALJs has increased over time. By contrast, when considering a fixed measure of outlierness, the proportion of ALJs with allowance rates in excess of 80 percent has declined significantly, from 31 percent in 2005 to 5 percent in 2014, while the proportion with allowance rates of less than 20 percent has increased slightly, from 0.3 percent to 1 percent. Overall, the distribution of ALJ allowance has become less skewed and more symmetrical over time.

The proportion of cases (not shown) decided by ALJs with approval rates in excess of 80 percent has dropped even more than the proportion of judges, from nearly 34 percent to 4.5 percent, because the number of cases decided by high-approval judges has declined proportionately over the 2005–2014 period. Among the low-approval judges, the change in the proportion of cases is not significantly different because low-approval judges have always decided a smaller number of cases. The aggregate number of cases has increased from about 430,000 in 2005 to about 520,000 in 2009 to 640,000 in 2013, while the distribution among judges has become more symmetrical around the mean from 2009 to 2013, although it still skews to the right.

Table 4 translates these indications of presumptively wrong decisions, in both directions, into dollar terms, on net, representing losses to taxpayers from the SSDI trust fund and the general fund of the Treasury (for SSI). The annual loss was more than \$10 billion in 2009, and declined to almost \$1.5 billion in 2014. Over the entire 2005–2014 period, the loss to taxpayers

has come to more than \$72 billion. These numbers are not discounted to present value and represent the value of future benefits. When resources and attention are devoted to improving program management, they are successful. At the same time, the structural and financial state of the disability program is dire, and while improvements are being made, the continuing high rate of presumptive mistakes adds billions to a growing deficit.

**Table 4. Annual Net Cost of Presumption Mistakes of High- and Low-Approval Administrative Law Judges (billions of dollars)**

Year	High approval	Low approval	Sum
2005	10.4	(1.2)	9.1
2006	11.3	(1.6)	9.8
2007	11.0	(1.5)	9.5
2008	10.1	(1.5)	8.6
2009	11.9	(1.6)	10.3
2010	10.7	(1.5)	9.2
2011	8.4	(1.3)	7.1
2012	5.1	(1.0)	4.1
2013	3.6	(0.7)	3.0
2014	2.4	(1.0)	1.4

Source: Authors' calculations.

Note: all numbers are rounded to the nearest tenth.

We next conduct a simple least squares regression analysis of the annual approval rates of ALJs, considering only judges with at least three years of experience. The underlying data come from the SSA. In particular, we want to see if judges' approval rates are related to the standard deviations of their own approval rates from 2005 through 2014, to the number of decisions they have made annually, and to the year of the decision. We represent the model in equation form as follows:

$$Y = B_0 + B_1X_1 + B_2X_2 + B_iX_t + u,$$



where  $Y$  is the judge's approval rate, 2005–2014;  $X_1$  is the standard deviation of the judge's approval rate, 2005–2014;  $X_2$  is the annual number of decisions made by the judge; and  $X_t$  are time dummy variables for 2005–2014. Table 5 shows the results.

**Table 5. Regression Analysis of Administrative Law Judge Approval Rates**

Source	SS	df	MS			
Model	68.9878261	11	6.27162055	Number of obs =	13219	
Residual	340.032724	13207	.025746401	F( 11, 13207) =	243.59	
				Prob > F =	0.0000	
				R-squared =	0.1687	
				Adj R-squared =	0.1680	
Total	409.02055	13218	.030944209	Root MSE =	.16046	

AllowanceRate	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
SDofAllowanceRate	-.4728226	.0296069	-15.97	0.000	-.5308564	-.4147889
Decisions	.0000523	8.00e-06	6.54	0.000	.0000367	.000068
year1	0	(omitted)				
year2	-.0072872	.006784	-1.07	0.283	-.0205848	.0060104
year3	-.0073604	.0067725	-1.09	0.277	-.0206354	.0059146
year4	-.0165027	.006867	-2.40	0.016	-.029963	-.0030425
year5	-.0039769	.0065499	-0.61	0.544	-.0168156	.0088618
year6	-.0380327	.0065085	-5.84	0.000	-.0507904	-.0252751
year7	-.0874753	.0063926	-13.68	0.000	-.1000057	-.0749449
year8	-.1370428	.0063076	-21.73	0.000	-.1494066	-.1246791
year9	-.1583018	.0063586	-24.90	0.000	-.1707656	-.1458379
year10	-.1687925	.006448	-26.18	0.000	-.1814314	-.1561535
_cons	.7264799	.0062145	116.90	0.000	.7142987	.7386612

Source: Authors' analysis based on SSA data.

Several interesting results obtain from this analysis. First, the higher the standard deviation of an ALJ's past decisions, or the more variability in an ALJ's decision-making, the lower the judge's current approval rate is. Stated another way, the more the judge decides in one direction consistently, the higher the judge's approval rate. This is a disconcerting finding because, given the randomness by which cases are supposed to be assigned to judges by the SSA, we should expect variability from year to year in the judge's approval rates. But high-approval judges are high-approval judges, year in and year out, whereas low-approval judges are not: they

are more variable over the years. We also see a positive coefficient on the number of cases decided—another concerning result. The more cases a judge decides, the higher the approval rate. This outcome should not occur in a well-functioning review system. Finally, we see significant and increasingly negative coefficients on the dummy variables for the last five years, as the economy has improved and as SSA administrative reforms, described later, have begun to have an influence.

### **Further Circumstantial Evidence**

The agency rules and their administration affect the overall award rates as well as the incentive to file an appeal. An interesting, related, and relevant issue is the December 2014 bankruptcy of Binder and Binder, the largest law firm specializing in SSDI claim appeals. This bankruptcy has been attributed to the tightening of administrative procedures lowering award rates and hence lowering payments to third-party representatives, especially attorneys. At the same time, the number of ALJ decisions fell noticeably in 2014.<sup>28</sup>

### **Recommendations**

SSA Commissioner Michael Astrue instituted several administrative reforms in response to the problems with the ALJ system. He hired and trained a record number of new ALJs, drawn from fresh candidate lists, even as some long-serving ALJs retired. He limited the number of cases that could be heard by any ALJ each year to 1,000, which was subsequently reduced to about 800. To limit claimant and representative abuses in the adjudication process, he allowed only one

---

<sup>28</sup> Sara Randazzo and Damian Paletta, “Social Security Disability Firm Binder & Binder Prepares for Possible Chapter 11: Company Faces Shrinking Demand for Services, as Government Scrutiny of Claims Tightens,” *Wall Street Journal*, December 12, 2014.

application for benefits per worker in the system at a time. He set up a more rigorous method of ensuring the rotation of cases among judges in response to clear signs of judicial collusion with attorneys. Additionally, he began a program of random review of ALJ decisions, including preeffectuation reviews of allowances.

Astrue's reforms were good and necessary but do not go far enough. Losses to taxpayers continue despite the recent reduction. Moreover, as public attention moves elsewhere or as demands to fix the claims backlog arise again, as indeed the backlog has again grown recently, his changes can easily be undone, because administrators face heavy political pressures to expedite and shortcut responsible processes. Future administrators could reverse reforms intentionally—or unintentionally, as bad habits slip back into the system. Astrue's program to increase accountability through random reviews of ALJ decisions and to increase judicial turnover by hiring new judges and encouraging a few to retire (prominently, Judge Daugherty) should therefore be made permanent and stronger. In particular, Congress should institute 15-year term limits for judges, who essentially enjoy lifetime tenure, to ensure that fresh legal minds are joining the pool of judges and to prevent it from becoming stale and unresponsive to legal criteria and requirements. A term of a decade and a half is long enough to insulate judges from administrative and political pressures and prevent undue political influence. Also, the conducting of a statistically valid number of preeffectuation reviews on ALJ allowances should not be at the agency's discretion; it should be required by statute, as it is at the DDS level.

The system faces a large backlog of cases, likely made worse by claimants who are strategic: for example, they can file serial claims in the hopes of eventually getting a lenient judgment at the initial or the appeals level. Congress can limit this gamesmanship by allowing only one application per claimant in a three-year period. This change would reduce the number

of claims. Those claims that were made would be more serious and substantive, thereby reducing the claims backlog. “On-the-record” decisions should not be allowed; these are difficult for the agency to subsequently review for program eligibility in continuing disability reviews because the documentation and hence delineation of the basis of the original determination is so poor. As we mentioned earlier, because judges must marshal more documentation for a denial than for an approval, they have an incentive to grant benefits to keep the system chugging along. The agency can fix this problem by further limiting the number of cases each judge must decide to 500: that is, about two and a quarter cases per working day, a reasonable number in view of the complexity of disability adjudication.

The system is further complicated because under the so-called “three hat” rule, the judge must advocate for the claimant, advocate for the government (that is, the taxpayer), and render unbiased judgement. Even if a claimant has legal counsel, the judge must still advocate on the claimant’s behalf. This rule must end. Most claimants—85 percent—now have third-party representatives, most of whom are experienced and are paid if they win the case. They can be expected to represent their clients well. Moreover, these professionals should be held responsible for getting supporting materials into court expeditiously and completely so the record can be closed in a timely manner before the hearing is held.

## **Conclusions**

In Lawrence Summers’s “The Scientific Illusion in Empirical Macroeconomics,” the celebrated academic articulates a standard of evidence often ignored in economics.<sup>29</sup> Summers brushes aside “statistical pyrotechnics,” noting that “physicists do not compete to find more and more

---

<sup>29</sup> Lawrence Summers, “The Scientific Illusion in Empirical Macroeconomics,” *Scandinavian Journal of Economics* 93, no. 2 (1991): 129–48.

elaborate ways to observe falling apples.” He advocates instead for an agenda based on “pragmatic economic work,” where “many different types of data are examined” and “no single test is held out as decisive.” Although strong results from one method can indicate a likely problem, examining multiple strands of administrative law judge data from the previous decade enables us to conclude definitively that there are systemic problems in the allowance process.

Using case studies of “outlier” judges, quality review data from the SSA, and empirical work pertaining to the iron triangle of dispositional volume, allowance rate, and decisional quality, we conclude that outlier high-allowance judges are deviating from the law in overproviding benefits. This tendency carries large economic stakes; removing both the most and least “generous” judges would have saved taxpayers more than \$72 billion in the 2005–14 period. But even this large amount fails to capture the tremendous opportunity cost that comes with having capable workers out of the labor force, a cost that will continue to rise as worker productivity grows over time. By reducing the number of applications, further capping the yearly number of cases heard by ALJs, ending lifetime ALJ tenure, devoting greater time and resources to conducting quality reviews, and ending judicial advocacy for claimants, we can restore sustainability and integrity to a troubled claims appeals process.